

Docket No.: 62807-016

PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE



Inventor Application of

Toshio OSHIMA, et al.

Serial No.: 09/978,183

Group Art Unit: 2184

Filed: October 17, 2001

Examiner:

For: A FAILURE SUPERVISING METHOD AND APPARATUS

RECEIVED  
DEC 05 2001  
Group 2100

TRANSMITTAL OF CERTIFIED PRIORITY DOCUMENT

Honorable Commissioner for Patents and Trademarks  
Washington, D. C. 20231

Sir:

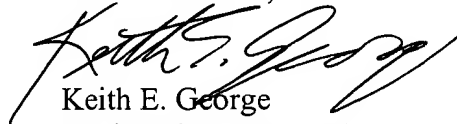
At the time the above application was filed, priority was claimed based on the following application:

**Japanese Patent Application No. 2001-045950, filed February 22, 2001**

A copy of each priority application listed above is enclosed.

Respectfully submitted,

MCDERMOTT, WILL & EMERY

  
Keith E. George  
Registration No. 34,111

600 13<sup>th</sup> Street, N.W.  
Washington, DC 20005-3096  
(202) 756-8000 KEG:mlw  
**Date: December 3, 2001**  
Facsimile: (202) 756-8087

62801-016  
Saeed OSHIMA et al.  
October 17, 2001  
09/978,183 #5  
McDermott, Will & Emery



日 本 国 特 許 庁  
JAPAN PATENT OFFICE

別紙添付書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出 願 年 月 日  
Date of Application:

2001年 2月22日

出 願 番 号  
Application Number:

特願2001-045950

出 願 人  
Applicant(s):

株式会社日立製作所

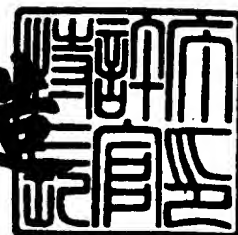
RECEIVED  
DEC 05 2001  
Group 2100

CERTIFIED COPY OF  
PRIORITY DOCUMENT

2001年 9月28日

特 許 庁 長 官  
Commissioner,  
Japan Patent Office

及 川 耕 造



【書類名】 特許願

【整理番号】 K01000941A

【提出日】 平成13年 2月22日

【あて先】 特許庁長官

【国際特許分類】 G06F 11/00

【発明者】

    【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所 システム開発研究所内

    【氏名】 大島 訓

【発明者】

    【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所 システム開発研究所内

    【氏名】 新井 利明

【発明者】

    【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所 システム開発研究所内

    【氏名】 佐藤 雅英

【発明者】

    【住所又は居所】 愛知県尾張旭市晴丘町池上 1 番地 株式会社日立製作所 情報機器事業部内

    【氏名】 鵜飼 宏樹

【特許出願人】

    【識別番号】 000005108

    【氏名又は名称】 株式会社日立製作所

【代理人】

    【識別番号】 100075096

    【弁理士】

    【氏名又は名称】 作田 康夫

【手数料の表示】

【予納台帳番号】 013088

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プールの可否】 要

【書類名】 明細書

【発明の名称】 障害監視方法及び装置

【特許請求の範囲】

【請求項 1】

タイマを用いてシステムの障害を監視する方法であって、

- 1) タイマを起動して前記タイマのリセットを判定し、
- 2) 前記タイマがリセットされていなければ、前記タイマのカウントダウンを実行し、
- 3) 所定の時間にタイムアウトしているかを判定し、
- 4) タイムアウトであれば、障害回復のための信号を発生し、
- 5) 障害から抜け出せなかった場合は、次のタイマに対して前記 1) から 4) のステップを繰り返し実行することを特徴とする障害監視方法。

【請求項 2】

請求項 1 の障害監視方法であって、前記ステップ 4) で発生する信号に応じて、フラグを立てる手順、割り込み信号を出力する手順、マスク不可能割り込みを出力する手順、及びシステムリセットを出力する手順を順次実行することによって、前記ステップ 5) の段階を経るごとに障害の程度に応じて障害回復を行なうことを特徴とする障害監視方法。

【請求項 3】

請求項 1 の障害監視方法であって、前記タイマをリセットする条件を複数個設定し、前記ステップ 5) の段階を経るごとに実施すべきタイマリセットと前記条件とを対応付けることを特徴とする障害監視方法。

【請求項 4】

請求項 1 及び請求項 2 記載の障害監視方法であって、前記ステップ 4) において、前記発生した信号に応じて実行された手順を記録することを特徴とする障害監視方法。

【請求項 5】

タイマを用いてシステムの障害を監視する装置は、

- 1) タイマを起動して前記タイマのリセットを判定する手段、

- 2) 前記タイマがリセットされていなければ、前記タイマのカウントダウンを実行する手段、
- 3) 所定の時間にタイムアウトしているかを判定する手段、
- 4) タイムアウトであれば、障害回復のための信号を発生する手段、
- 5) 障害から抜け出せなかった場合は、次のタイマに対して前記1) から4) のステップを繰り返し実行する手段を有することを特徴とする障害監視装置。

【請求項6】

請求項5の障害監視装置であって、前記信号を発生する手段からの信号に応じて、フラグを立てる手順、割り込み信号を出力する手順、マスク不可能割り込みを出力する手順、及びシステムリセットを出力する手順を順次実行することによって、前記繰り返し実行する手段による段階を経るごとに障害の程度に応じて障害回復を行なうことを特徴とする障害監視装置。

【請求項7】

請求項5の障害監視装置であって、前記タイマをリセットする条件を複数個設定し、前記繰り返し実行する手段による段階を経るごとに実施すべきタイマリセットと前記条件とを対応付けることを特徴とする障害監視装置。

【請求項8】

請求項5及び請求項6記載の障害監視装置であって、前記信号の発生手段は、前記発生した信号に応じて実行された手順を記録する手段を有することを特徴とする障害監視装置。

【請求項9】

タイマを用いてシステムの障害を監視する方法であって、

- 1) 起動したタイマのリセットされていなければ、前記タイマのカウントダウンを実行し、
- 2) 所定の時間にタイムアウトしていれば、障害回復のための手順を実行し、
- 3) 障害から抜け出せなかった場合は、次のタイマに対して前記1) 及び2) のステップを繰り返し実行して、段階を経るごとに障害の程度に応じて障害回復を行なうことを特徴とする障害監視方法。

【発明の詳細な説明】

【 0 0 0 1 】

【発明の属する技術分野】

本発明はシステムの障害監視に関し、特に拡張機器からの割り込みによる計算機システムの障害監視に関する。

【 0 0 0 2 】

【従来の技術】

システムの障害監視方法として、ウォッチドッグタイマ(以下WDT)と呼ばれる方法がある。WDTとは、タイマによって経過時間を計測し一定時間経過するとシステムを再起動する仕組みであり、システムは正常に機能している間、このタイマを定期的のリセットすることにより、システム再起動が実行されるのを妨げる。システムが暴走状態に陥り、WDTをリセットすることが出来なかった場合、タイマはタイムアップし、システム全体を再起動する。この仕組みにより、システムは動作を継続することができる。

またWDTの関連技術として、タイマタイムアウト後、フラグをセットするものや通常の割り込み、マスク不可能割り込み(NMI)を生成するものもある。

【 0 0 0 3 】

【発明が解決しようとする課題】

システム管理者は、システムに障害が発生した場合、可能な限りサービスを停止せずに回復させたいと考えている。また、障害停止による再起動が避けられないとしても、可能な限り障害時の情報を収集することで、障害の再発を防止したいと考えている。

【 0 0 0 4 】

しかし、単純なWDTはシステムが暴走した際に、システムを再起動するだけである。障害の種類によっては、例えばシステムに割り込みを発生させて、障害から復帰したり障害時の情報を収集することによって、障害の再発防止に役立てることができる。また、WDTタイムアウト後、システムに割り込みをかけるだけのWDTでは、割り込みによっては、障害から復帰できないような深刻な状態に陥ってしまった場合、システムが停止してしまうといった問題がある。

【 0 0 0 5 】

さらに、従来のWDTはタイマをリセットする方法として、タイマリセットポートにリセット用データをセットするかまたは、WDTリセット命令を出力するといった方法を提供してきた。しかし、システム内に複数の処理装置が存在するようなケースで、その中の一つの処理装置に発生した障害だけでも検出したいという場合、従来の方法ではそれを実現することができなかった。

#### 【0006】

また、障害回復の手法には割り込み、NMI、システムリセットがあるが、これらにはそれぞれ以下のような長所と欠点がある。

#### 【0007】

すなわち、割り込みによる障害回復は、クリティカルリージョンを壊してしまいシステムの継続が困難であり、さらに、不揮発性記憶装置に記録されていないシステムの状態をリセットしてシステムを再起動しなくても障害から回復できるが、割り込み禁止状態、あるいは、割り込みを受け取れてもシステムが動作できない状態にあった場合、障害回復に対応することが出来ない。

#### 【0008】

NMIによる障害回復は、割り込み禁止状態に陥った場合に対応できず、さらに、不揮発性記憶装置に記録されていないシステムの状態をリセットしてシステムを再起動しなくても障害から回復できるが、クリティカルリージョンを侵している可能性を否定できないので、システム安定のため一旦システム再起動を入れる必要が発生するという問題がある。

#### 【0009】

システムリセットによる障害回復は、これにより対応できないシステム状態は無いが、不揮発性記憶装置に記憶されていない情報を全てリセットしてしまうため、管理者には障害時のシステム状態がわからず、障害再発防止の措置をとるための情報に乏しいという問題がある。

#### 【0010】

##### 【課題を解決するための手段】

本発明のWDTは複数段階のWDTが、段階を経るごとにシステムに対しより強力に介入する割り込みを出力する。即ち、本発明では、障害の程度に応じて割り込み



の種類（程度）を変え、その割り込みに応じて障害回復処理を行なえるようにする。

#### 【0011】

例えばタイマの第1段階がタイムアウトした場合、システムに割り込みを入れると共に第2段階のWDTを起動する。第1段階の割り込みによってシステムが障害から抜け出すことができた場合、システムはWDTをリセットするかまたは停止するといった処置をとるものとする。第1段階の割り込みで障害から抜け出すことができなかった場合は、第2段階のWDTがタイムアウトして割り込み、またはマスク不可能割り込み(以下NMI)等を入力する。この割り込みによって障害から抜け出すことができなかった場合、更に第3段階のWDTが起動される。第3段階のWDTがタイムアウトした場合、システムリセットを入れることによって、システムを再起動する。

#### 【0012】

また、WDTをリセットする手段として、複数のポートを持ったWDTリセットポートを用意する。このような仕組みを提供することで、マルチプロセッサシステムのような複数の処理装置が並列に処理をしているシステムにおいて、処理装置のうちひとつが停止してしまったような障害を検出できる。

#### 【0013】

##### 【発明の実施の形態】

次に、本発明について図面を用いて詳細に説明する。

#### 【0014】

図1は、本発明の一実施例に係る障害監視装置の動作を示すフローチャートと障害監視装置を制御するレジスタの構成を示すブロック図を併せたものである。図2は、図1の不揮発性メモリ124の内部を示したものである。ステップ101から117は、3段階のウォッチドッグタイマWDTの動作を示している。

#### 【0015】

本障害監視装置では、ステップ101から起動されると、まずWDT1を起動(ステップ102)する。次にWDT1がリセットされたかを検査(ステップ103)する。WDTのリセット方法の詳細については後述する。もしWDT1がリセットされていればステップ

102に戻りWDT1を再起動する。ここでWDT1がリセットされていなければ、WDT1のカウントを進め(ステップ104)、WDT1がタイムアウトしたか否かを判定(ステップ105)する。この判定にはWDT1のタイムアウト時間121を設定値として用いる。WDT1がタイムアウトしていなければ、WDT1リセット判定103まで戻る。WDT1がタイムアウトしていれば、システムに割り込み信号を出力し、同時に不揮発性メモリ124内のWDT1タイムアウト201に、割り込み信号を出力したことを示す情報を出力し、WDT2を起動(ステップ107)する。

#### 【 0 0 1 6 】

WDT2もWDT1と同様にWDT2のリセット判定(ステップ108)を行い、WDT2のカウントダウン(ステップ109)をし、WDT2タイムアウト時間122を用いてWDT2のタイムアウト判定(110)をする。WDT2がリセットされた場合、フローはWDT1起動(ステップ102)まで戻る。WDT2がタイムアウトした場合、マスク不可割り込みNMI信号を出力し、NMI信号を出力したことを示す情報を不揮発性メモリ124のWDT2タイムアウト202に出力する(ステップ111)。そしてWDT3を起動する(ステップ112)。

#### 【 0 0 1 7 】

WDT3もWDT1,2と同様に動作をする。WDT3がタイムアウトした場合、リセット信号を出力したことを示す情報を不揮発性メモリ124のWDT3タイムアウト203に出力して、システムリセット信号を出力する。これによりシステム全体が再起動される。

#### 【 0 0 1 8 】

次にWDT1,2,3のリセット方法を述べる。WDTのリセットポート118は図のように8つのポートによって構成される。リセットポートの各ポートには、監視対象者(後述するOSなど)によってステータスなどの情報が定期的書き込まれる。各ポートはそれぞれ状態レジスタ119に対応するビットを持ち、各ポートにデータがセットされると状態レジスタ119の対応するビットがセットされる仕組みとなっている。障害監視装置は状態レジスタ119と予め設定された設定レジスタ120との比較を行い、値が一致した場合、状態レジスタ119をクリアしてWDTをリセットする。これはWDT1,2,3ともに同様の仕組みである。

## 【 0 0 1 9 】

また、ユーザ領域204は計算機システム上の上位ソフトウェアが使用するために解放された領域である。

## 【 0 0 2 0 】

図3は、図1に示した障害監視装置305を搭載し、特開平11-149385にあるようなマルチOS構成方法を用いて一つの処理装置を持つ一台の計算機上303で2つのオペレーティングシステムを動作させたものである。ここで第1OS301は通常の業務が行われるOSであり、このOS上では業務用のアプリケーションプログラムが動作している。これに対し第2OS304は第1OS301の生死をマルチ構成部302を通して監視している。もし、第1OS301が障害状態に陥ってしまったことを第2OS304が検出した場合、マルチOS構成部302の機能によって第1OSの状態を取得したり第1OSだけを再起動することによって、障害を回復することができる。さらに第2OS304は障害監視装置305を制御するデバイスドライバを持っており、起動時には障害監視装置305のWDTタイムアウト時間121,122,123を設定する。さらに設定レジスタ120にはリセットポート118のRST0に対応するビットだけをセットする。そして、第2OSはWDT1タイムアウト時間の範囲内で、定期的に障害監視装置305に対し、リセットポート118のRST0に情報を出力するという形で、自分が生存していることを示す生存信号を発信する。もし、第1OSや第2OSの障害によって第2OSが停止してしまった場合、生存信号の出力すなわちリセットポート118のRST0への信号出力も途絶え、それによってWDT1さらにはWDT2がタイムアウトし、マルチOS構成部302を通して第2OS304に割り込みやNMIが出力される。

## 【 0 0 2 1 】

通常、割り込みまたはNMIによって第2OS304は障害状態から復帰することができる。第2OS304の障害監視装置デバイスドライバは、WDTを停止させ障害情報の収集を開始する。まず、第2OSは障害監視装置305の図2に示す不揮発性メモリ124内のWDT1タイムアウト201またはWDT2タイムアウト202を参照することで、障害の程度を把握できる。出力されたものが割り込みであった場合、障害原因が第2OS304でなければ第1OS301の障害情報を第2OS304内に取得した上で、第1OS301だけを再起動することにより障害回復できる。

もし、障害原因が第2 OS304であったり、出力されたものが割り込みでなくNMI信号であった場合、第1 OS301や第2 OS304またはマルチOS構成部302のクリティカルリージョンが侵されている可能性が考えられるので、第2 OS304は第1 OS301から障害情報を収集し、不揮発性メモリ124内のユーザ領域204にその情報を記録したのち、システムリセットを発行しシステムを再立ち上げる。システム管理者は、再起動の後、このユーザ領域204に残った障害情報を取得することによって障害の再発防止対策を講じる手がかりとすることができる。

#### 【0022】

もし障害監視装置305が発生させた割り込みやNMIでも第2 OS304が障害から復帰できないような障害に陥ってしまった場合でも、システムはWDT3タイムアウト後のリセットにより再起動することで、障害停止だけは避けることができる。

#### 【0023】

図4は、図1に示した障害監視装置305を8つの処理装置401(以下CPU)と割り込み制御装置402をもつ計算機に搭載した例である。この計算機においては、割り込み制御装置が割り込みをどのプロセッサに伝達するか、あるいはマスク可能割り込みとして伝達するかどうかを決定することができる。計算機上のOSは障害監視装置用デバイスドライバを持っており、このデバイスドライバが障害監視装置305内の設定レジスタ120の全てのビットをセットすることで、リセットポート118の全てのポートを有効にする。CPU0はRST0、CPU1はRST1というように、各CPUはそれぞれ障害監視装置内の対応するリセットポートRST0～RST7に情報を出力することで、正常動作していることを障害監視装置に通知する。

#### 【0024】

もし、処理装置CPU0からCPU7のうちどれかひとつにでも障害が発生した場合、すべてのリセットポートRST0～RST7が書き換えられないため、状態レジスタ119と設定レジスタ120が一致しないので、WDTがリセットされず、従ってWDTがタイムアウトする。

#### 【0025】

WDTがタイムアウトすると障害監視装置305は、割り込み制御装置402を通して、各処理装置CPU0～CPU7に割り込みを入れるが、割り込みを入れる処理装置や割

り込みがマスク可能か不可能かといった設定は、割り込み制御装置402によって選択することが可能である。

【 0 0 2 6 】

【発明の効果】

以上に述べたように、本発明の障害回復方法は、複数段階のWDTを連動させる手順と連動したWDTが、段階的に強力にシステムに介入する手順と、割り込みによって回復できる障害には割り込みによって、マスク不可能割り込みによってしか回復できない障害にはマスク不可能割り込みにはマスク不可能割り込みによって、そして、システムリセットでしか回復できないような障害に対しては、システムリセットによって、障害に対応することができる。また、設定によって有効、無効を決定できる複数のポートをもつWDTリセットポートを用意することで、複数の処理装置が並列に動作するような計算機の障害をも監視できる。

【図面の簡単な説明】

【図 1】

本発明の一実施例である障害監視装置の動作を示すフローチャートと障害監視装置を制御するためのポートの構成を示すブロック図である。

【図 2】

図 1 中の不揮発性メモリの内部構成を示すブロック図である。

【図 3】

本発明の一実施例に係る計算機上のOSと障害監視装置の関係を示すブロック図である。

【図 4】

本発明の一実施例に係る複数の処理装置を搭載した計算機と障害監視装置の関係を示すブロック図である。

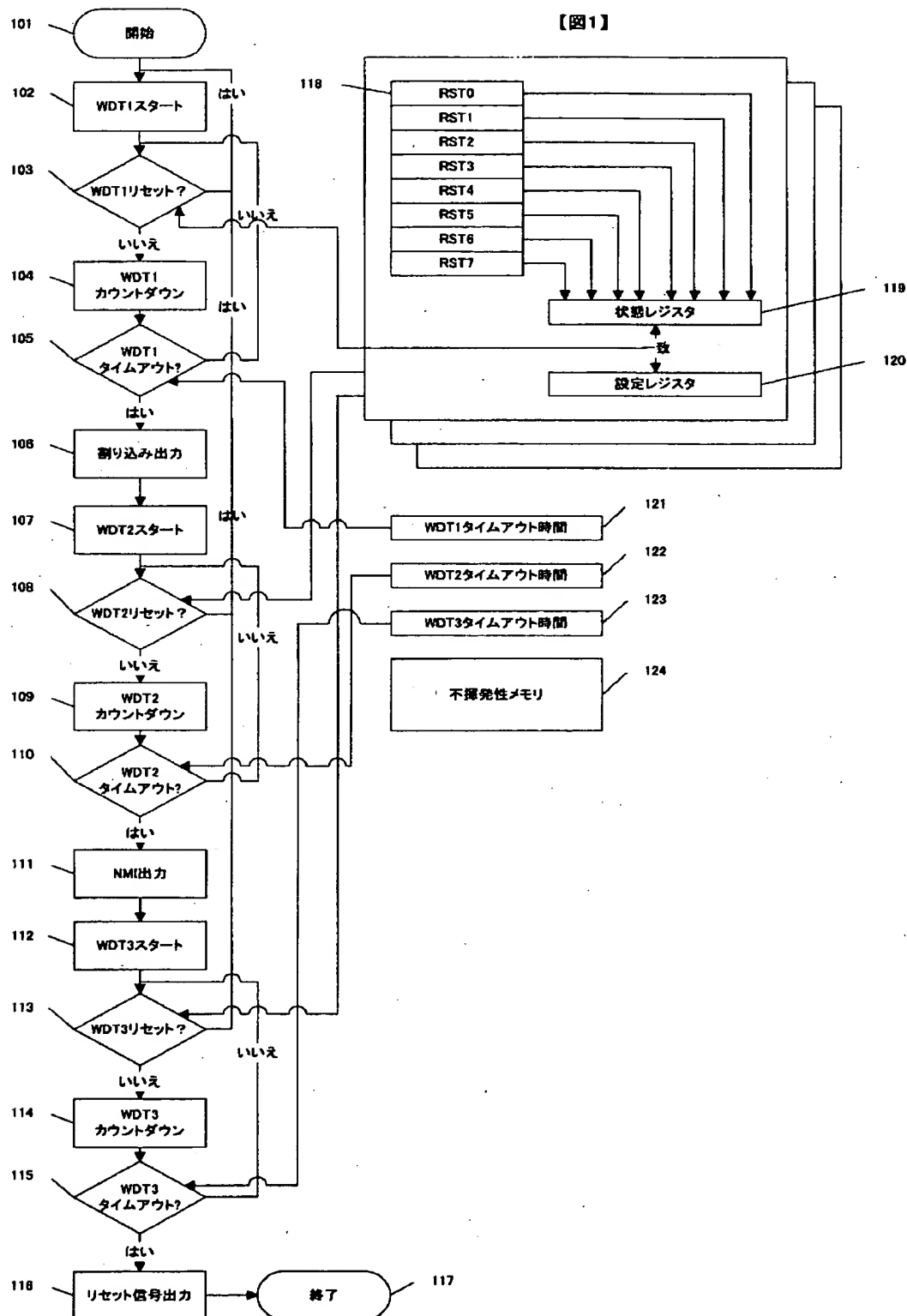
【符号の説明】

- 1 1 8 リセットポート
- 1 1 9 状態レジスタ
- 1 2 0 設定レジスタ
- 1 2 1 ウォッチドッグタイマ 1 タイムアウト時間

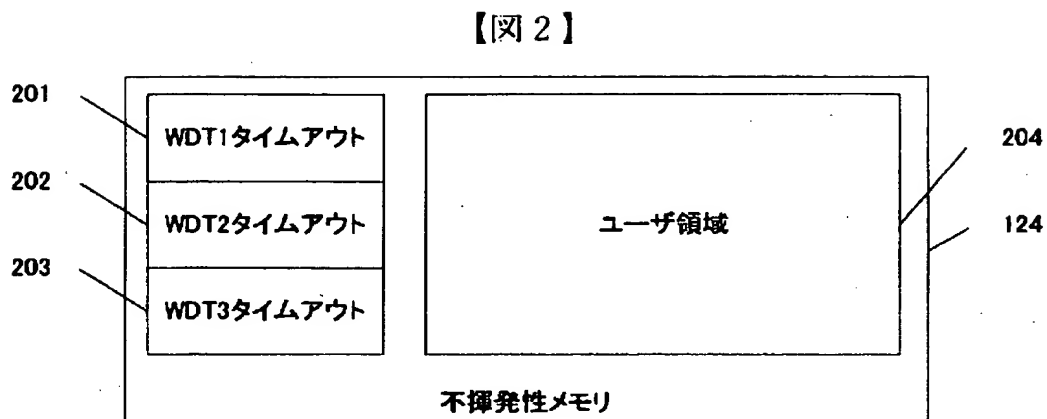
- 1 2 2 ウォッチドッグタイマ 2 タイムアウト時間
- 1 2 3 ウォッチドッグタイマ 3 タイムアウト時間
- 1 2 4 不揮発性メモリ
- 2 0 1 ウォッチドッグタイマ 1 タイムアウト
- 2 0 2 ウォッチドッグタイマ 2 タイムアウト
- 2 0 3 ウォッチドッグタイマ 3 タイムアウト
- 2 0 4 ユーザ領域
- 3 0 1 第 1 O S
- 3 0 2 マルチ O S 構成部
- 3 0 3 計算機
- 3 0 4 第 2 O S
- 3 0 5 障害監視装置
- 4 0 1 処理装置
- 4 0 2 割り込み制御装置

【書類名】 図面

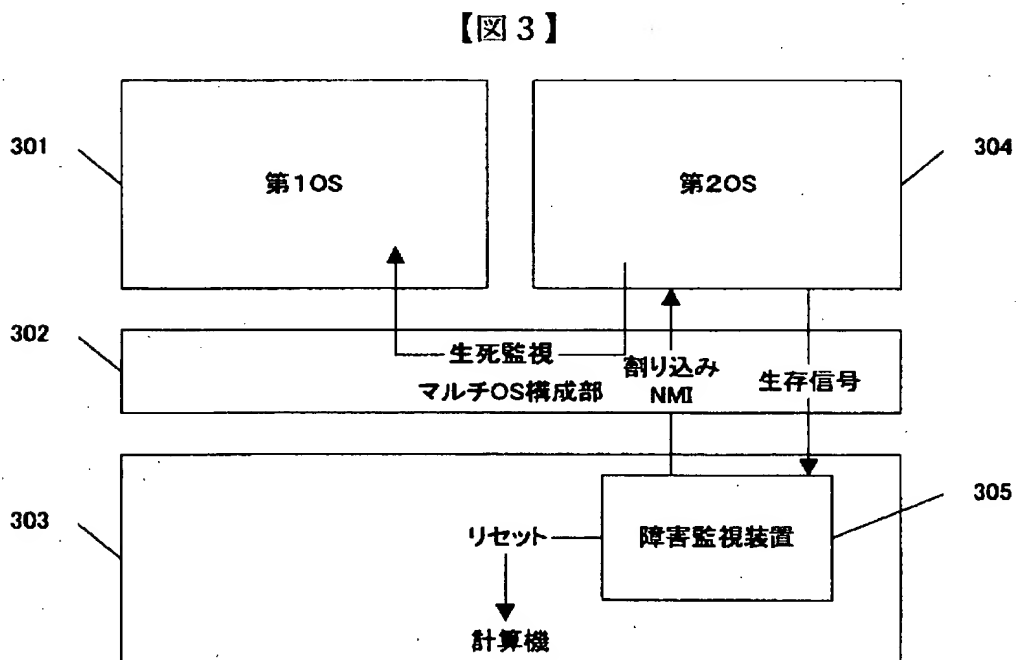
【図1】



【図 2】

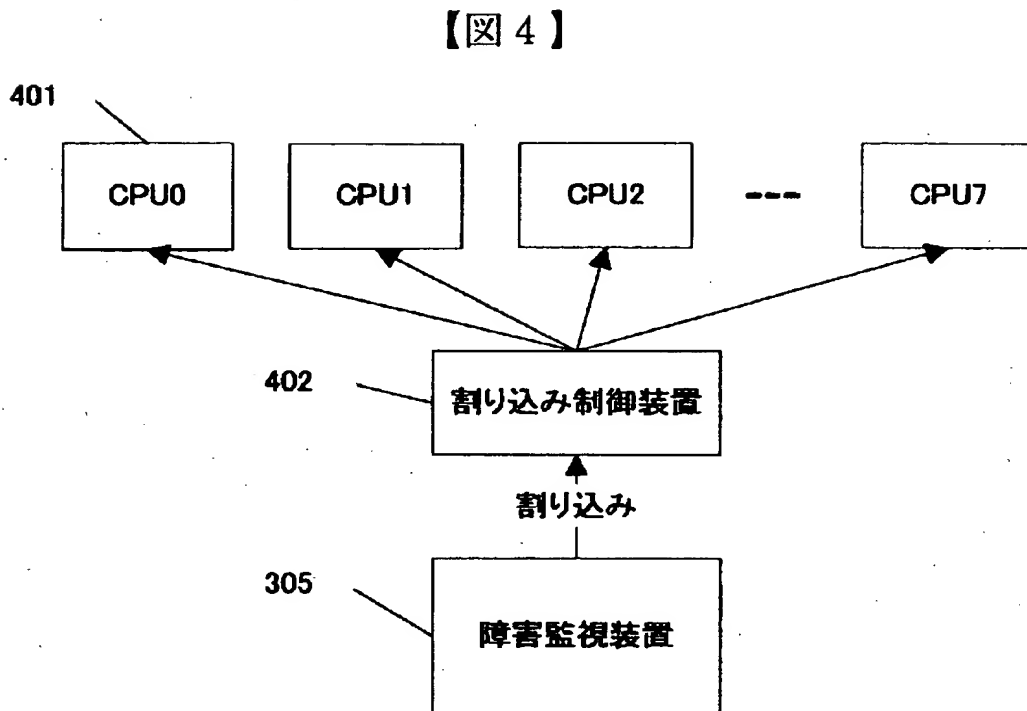


【図 3】





【図 4】



【書類名】 要約書

【要約】

【課題】 WDTタイムアウト後、システムに割り込みをかけるだけのWDTでは、割り込みだけでは障害から復帰できない深刻な状態に陥った場合、システムが停止する。

【解決手段】 複数段階のWDTを連動させる手順と連動したWDTが段階的に強ちにシステムに介入する手順と、割り込みによって回復できる軽度の障害には割り込みによって、マスク不可能割り込み以外では回復できない中程度の障害にはマスク不可能割り込みによって、そして、再起動以外に回復できないような重度の障害に対しては、システムリセットによって障害に対応する。また、設定によって有効、無効を決定できる複数のポートをもつWDTリセットポートにより、複数の処理装置が並列に動作する計算機の障害も監視する。

【選択図】 図 1

出 願 人 履 歴 情 報

識別番号 [000005108]

1. 変更年月日 1990年 8月31日

[変更理由] 新規登録

住 所 東京都千代田区神田駿河台4丁目6番地  
氏 名 株式会社日立製作所